Big Bang, Big Data, Big Iron

High Performance Computing and the Cosmic Microwave Background

Julian Borrill

Computational Cosmology Center, LBL

Space Sciences Laboratory, UCB

and the

BOOMERanG, MAXIMA, Planck, EBEX & PolarBear collaborations

Outline & Warning

1. A brief history of cosmology and the CMB
2. CMB physics and observations
3. CMB data analysis and high performance computing

Cosmologists are often in error

but *never* in doubt.

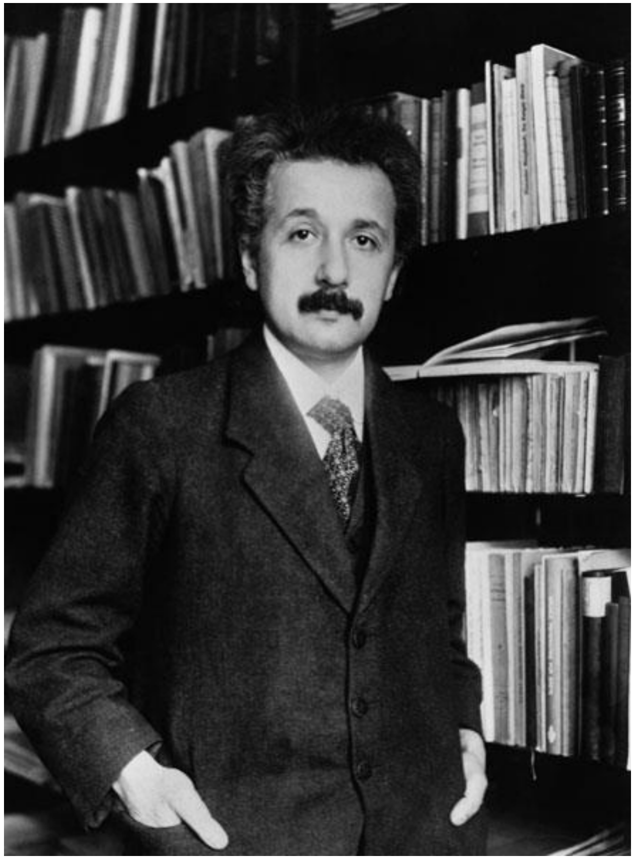
1916 – General Relativity

- General Relativity
  - Space tells matter how to move.
  - Matter tells space how to bend.

$$G_{\mu\nu} = 8 \pi G T_{\mu\nu}$$

Space *Matter*

- But this implies that the Universe is dynamic, and everyone *knows* it's static …
- … so Einstein adds a Cosmological Constant (even though the result is unstable equilibrium)



1929 – Expanding Universe

- Using the Mount Wilson 100-inch telescope Hubble measures nearby galaxies'
  - velocity (via their redshift)
  - distance (via their Cepheid variables)

  and finds

$$\boxed{v \propto d}$$

- Space is expanding!
- The Universe is dynamic after all.
- Einstein calls the Cosmological Constant "my biggest blunder".

1930-60s – Steady State vs Big Bang

- What does an expanding Universe tells us about its origin and fate?

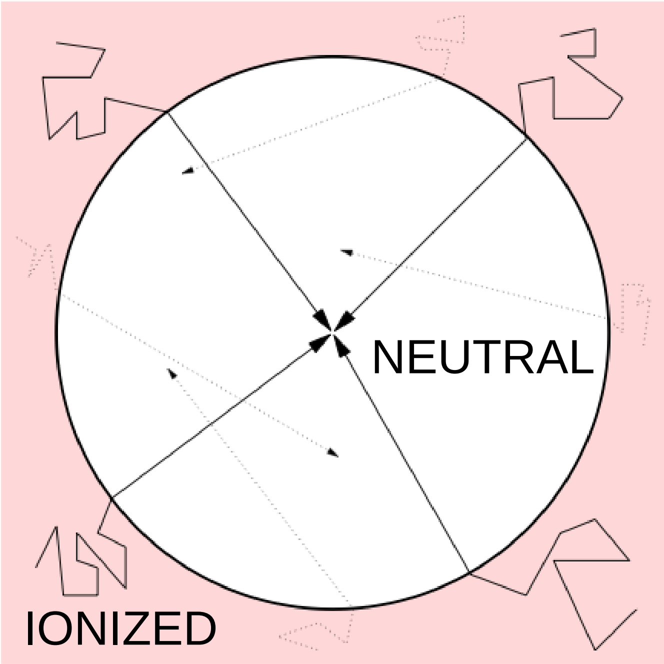
  – Steady State Theory:
    - new matter is generated to fill the space created by the expansion, and the Universe as a whole is unchanged and eternal (past & future).

  – Big Bang Theory:
    - the Universe (matter and energy; space and time) is created in a single explosive event, resulting in an expanding and hence cooling & rarifying Universe.

1948 – Cosmic Microwave Background

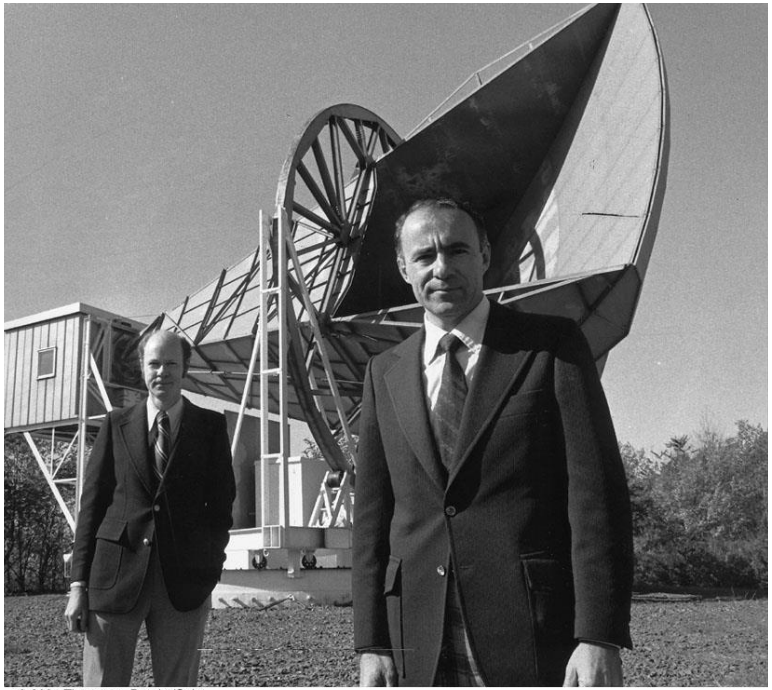
- In a Big Bang Universe the expanding Universe eventually cools through the ionization temperature of hydrogen: $p^+ + e^- \Rightarrow H$.
- Without free electrons to scatter off, the photons free-stream to us today.
- Alpher, Herman & Gamow predict a residual photon field at 5 – 50K

- COSMIC – filling all of space.
- MICROWAVE – redshifted by the expansion of the Universe from 3000K to 3K.
- BACKGROUND – primordial photons coming from “behind” all astrophysical sources.



1964 – First Detection

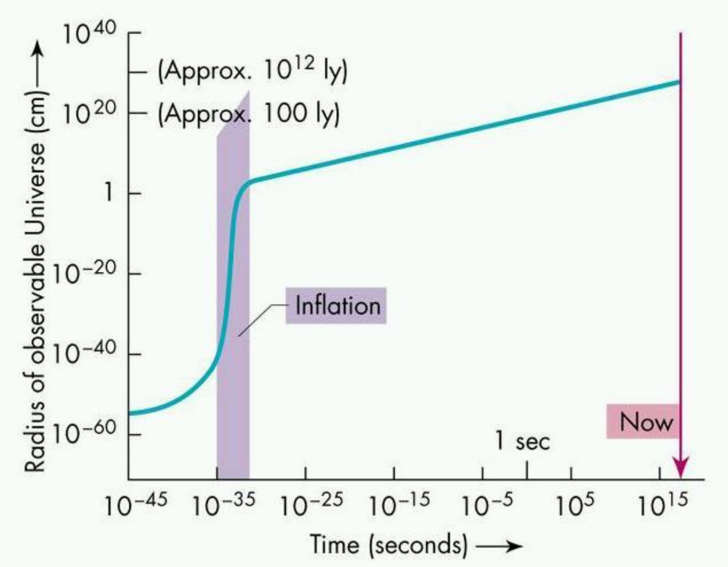
- While trying to zero a Bell Labs radio telescope, Penzias & Wilson found a puzzling residual signal that was constant in time and direction.
- They determined it wasn't terrestrial, instrumental, or due to a "white dielectric substance", but didn't know what it was.
- Meanwhile Dicke, Peebles, Roll & Wilkinson were trying to build just such a telescope in order to detect this signal.
- Penzias & Wilson's accidental measurement killed the Steady State theory and won them the 1978 Nobel Prize in physics.



1980 – Inflation

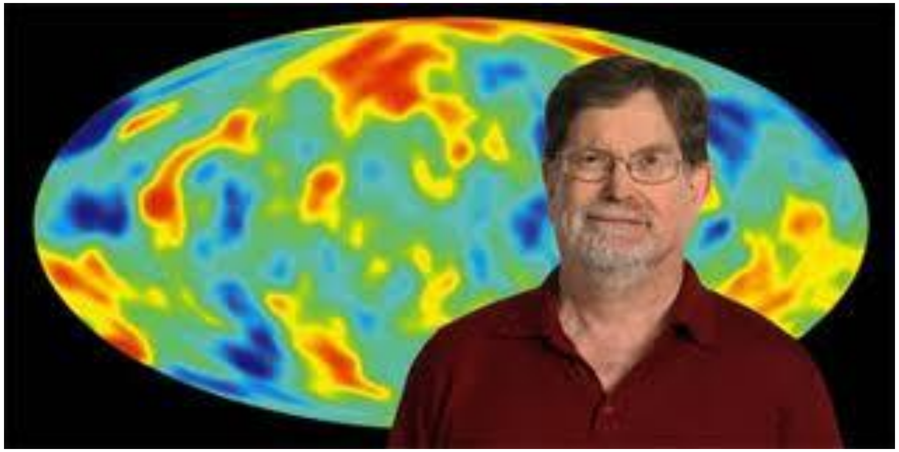
- More and more detailed measurements of the CMB temperature showed it to be uniform to better than 1 part in 100,000.
- At the time of last-scattering any points more than 1º apart on the sky today were out of causal contact, so how could they have exactly the same temperature? This is the horizon problem.

- Guth proposed a very early epoch of exponential expansion driven by the energy of the vacuum.
- This also solved the flatness & monopole problems.



1992 – CMB Fluctuations

- For structure to exist in the Universe today there must have been seed density perturbations in the early Universe.
- Despite its apparent uniformity, the CMB must therefore carry the imprint of these fluctuations.
- After 20 years of searching, fluctuations in the CMB temperature were finally detected by the COBE satellite mission.
- COBE also confirmed that the CMB had a perfect black body spectrum, as a residue of the Big Bang would.

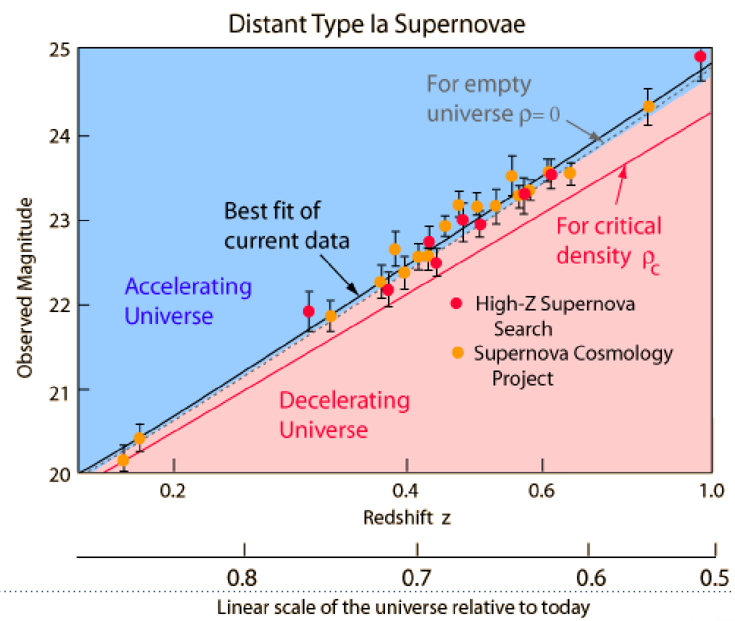


- Mather & Smoot share the 2006 Nobel Prize in physics.

1998 – The Accelerating Universe

- The fate and geometry of the Universe were thought to depend solely on the amount of matter it contained:
  - Below the critical density: eternal expansion, open Universe.
  - At critical density: expansion asymptotes to zero, flat Universe.
  - Above critical density: expansion turns to contraction, closed Universe.

- Measurements of the brightness and distances of supernovae surprisingly show the Universe is accelerating!
- Acceleration (maybe) driven by a Cosmological Constant!
- Perlmutter and Riess & Schmidt share 2011 Nobel Prize in physics.



2000 – The Concordance Cosmology

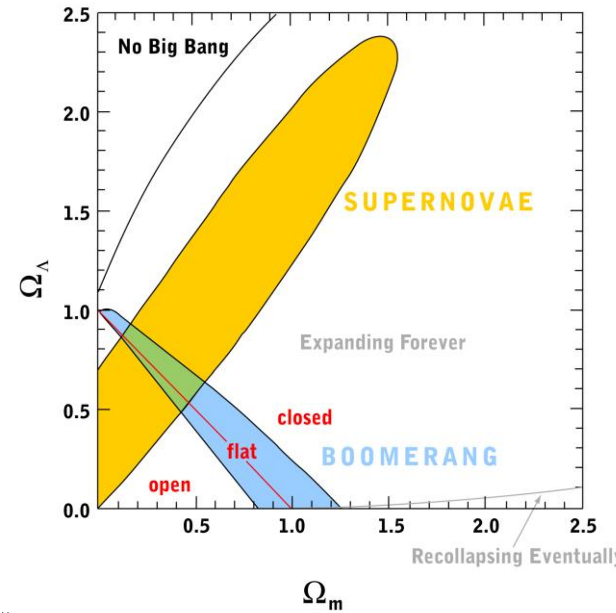
- The BOOMERanG & MAXIMA balloon experiments measure small-scale CMB fluctuations, demonstrating that the Universe is flat.
- The CMB fluctuations encode cosmic geometry ($\Omega_\Lambda + \Omega_m$)
- Type 1a supernovae encode cosmic dynamics ($\Omega_\Lambda - \Omega_m$)
- Their combination breaks the degeneracy in each.

The Concordance Cosmology:

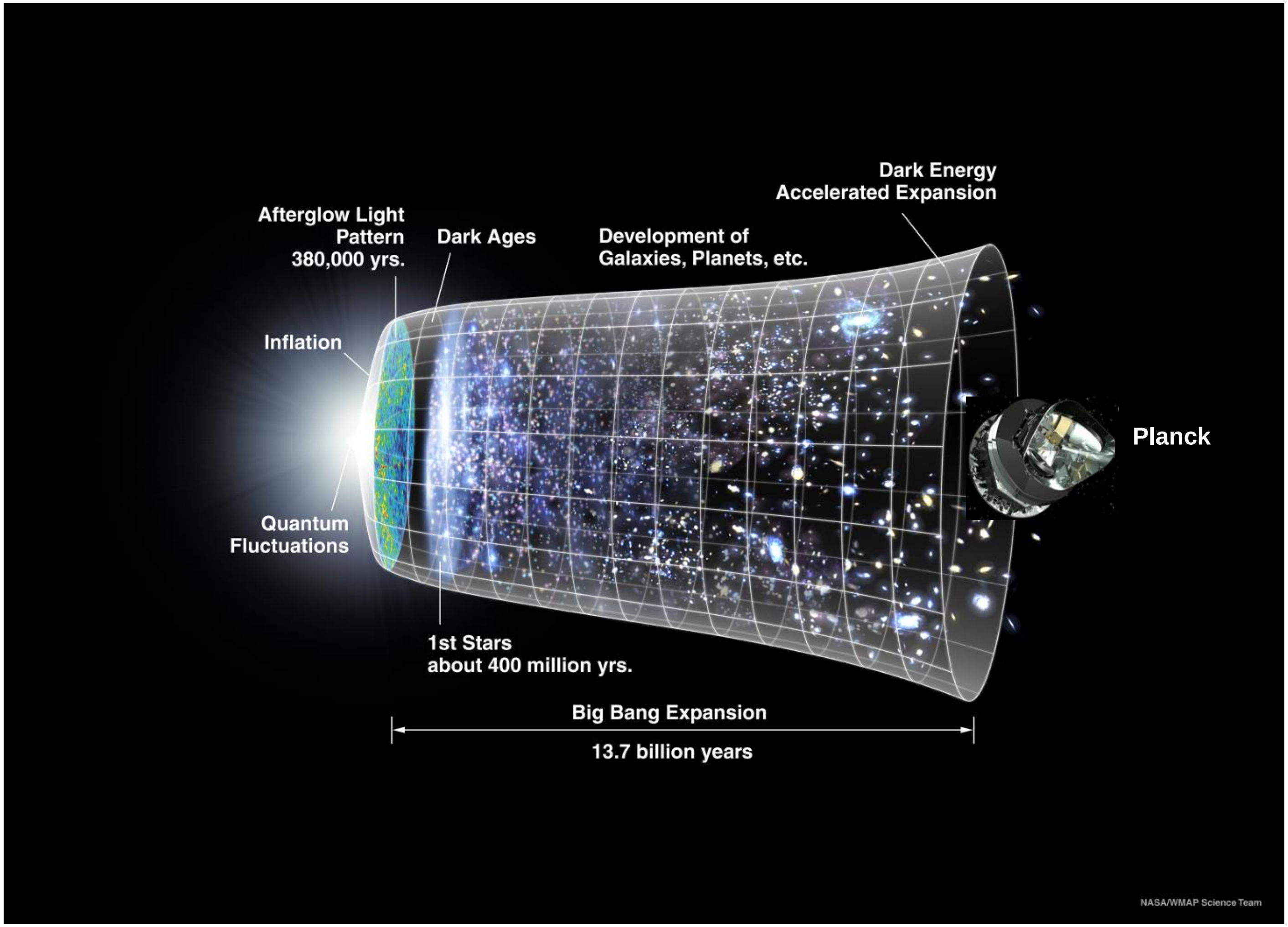
• 70% Dark Energy + 25% Dark Matter + 5% Baryons

=> 95% ignorance!

• What and why is the Dark Universe?



A History Of The Universe

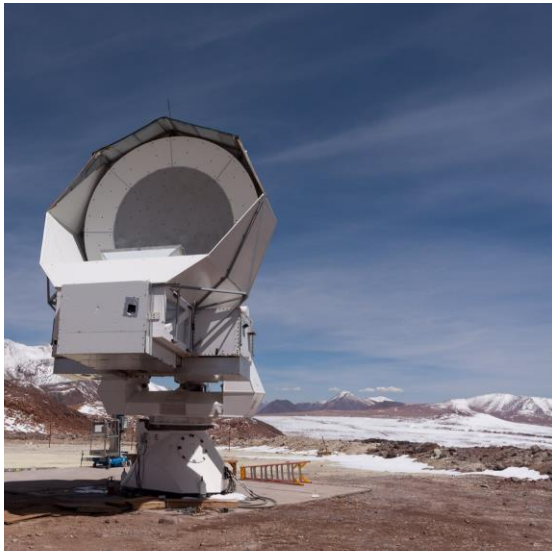


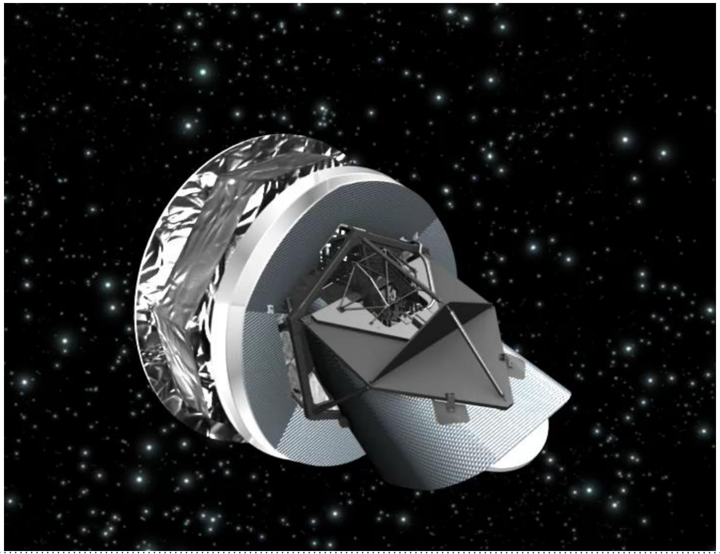
CMB Science

- Primordial photons trace the entire history of the Universe.
- Primary anisotropies:
  - Generated before last-scattering, encode all physics of the early Universe
    - Fundamental parameters of cosmology
    - Quantum fluctuation generated density perturbations
    - Gravity waves from Inflation
- Secondary anisotropies:
  - Generated after last-scattering, encode all physics of the later Universe
    - Gravitational lensing by dark matter
    - Spectral shifting by hot ionized gas
    - Red/blue shifting by evolving potential wells
- A repeating history of theoretical curiosity becoming observed signal.
- The challenges are (i) detection and (ii) decoding.

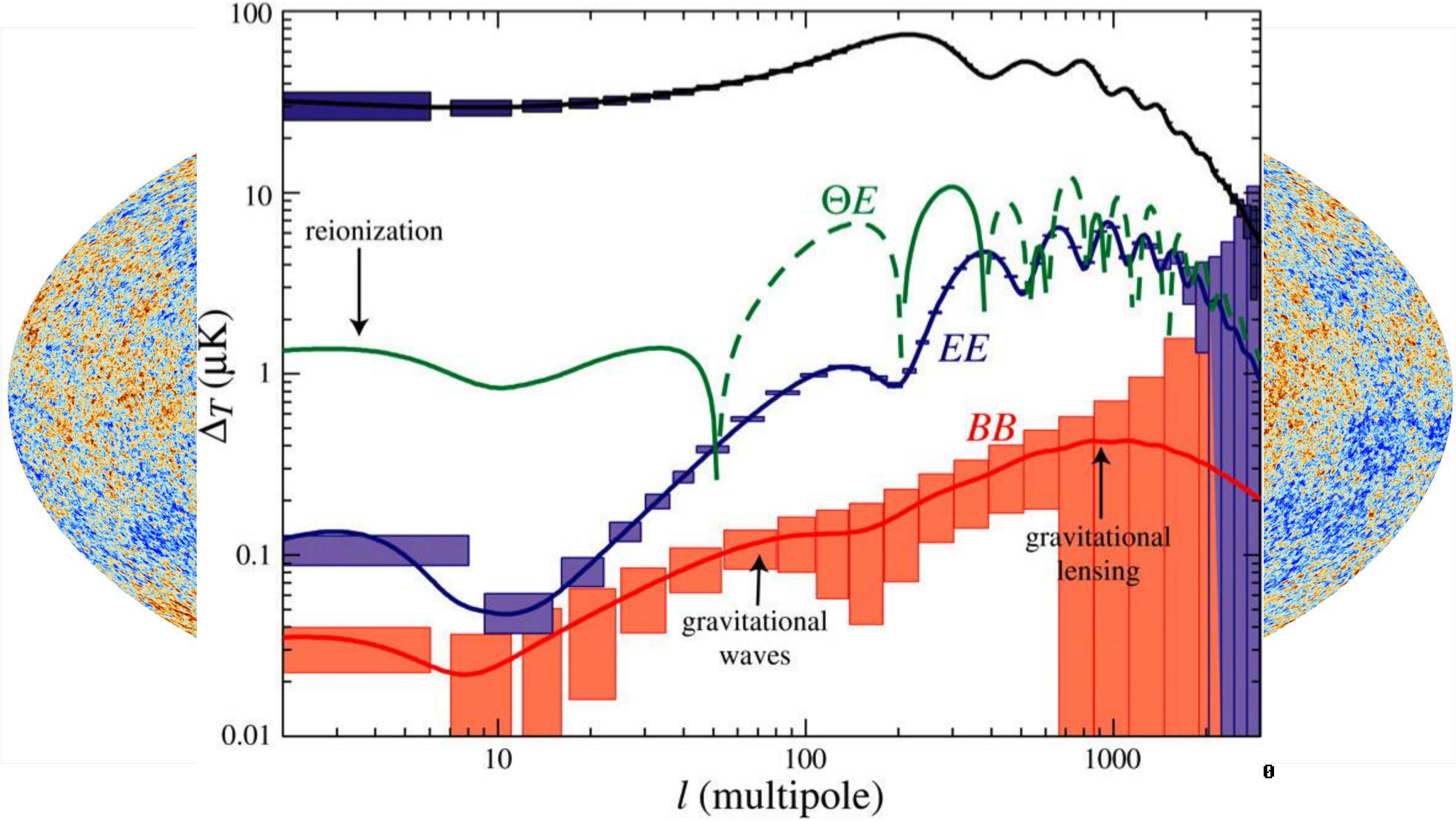
Detecting the CMB

- Searching for microK – nanoK fluctuations on a 3K background
- Need very many, very sensitive, very cold, detectors.
- Scan part of the sky from high dry ground or the stratosphere, or all of the sky from space.



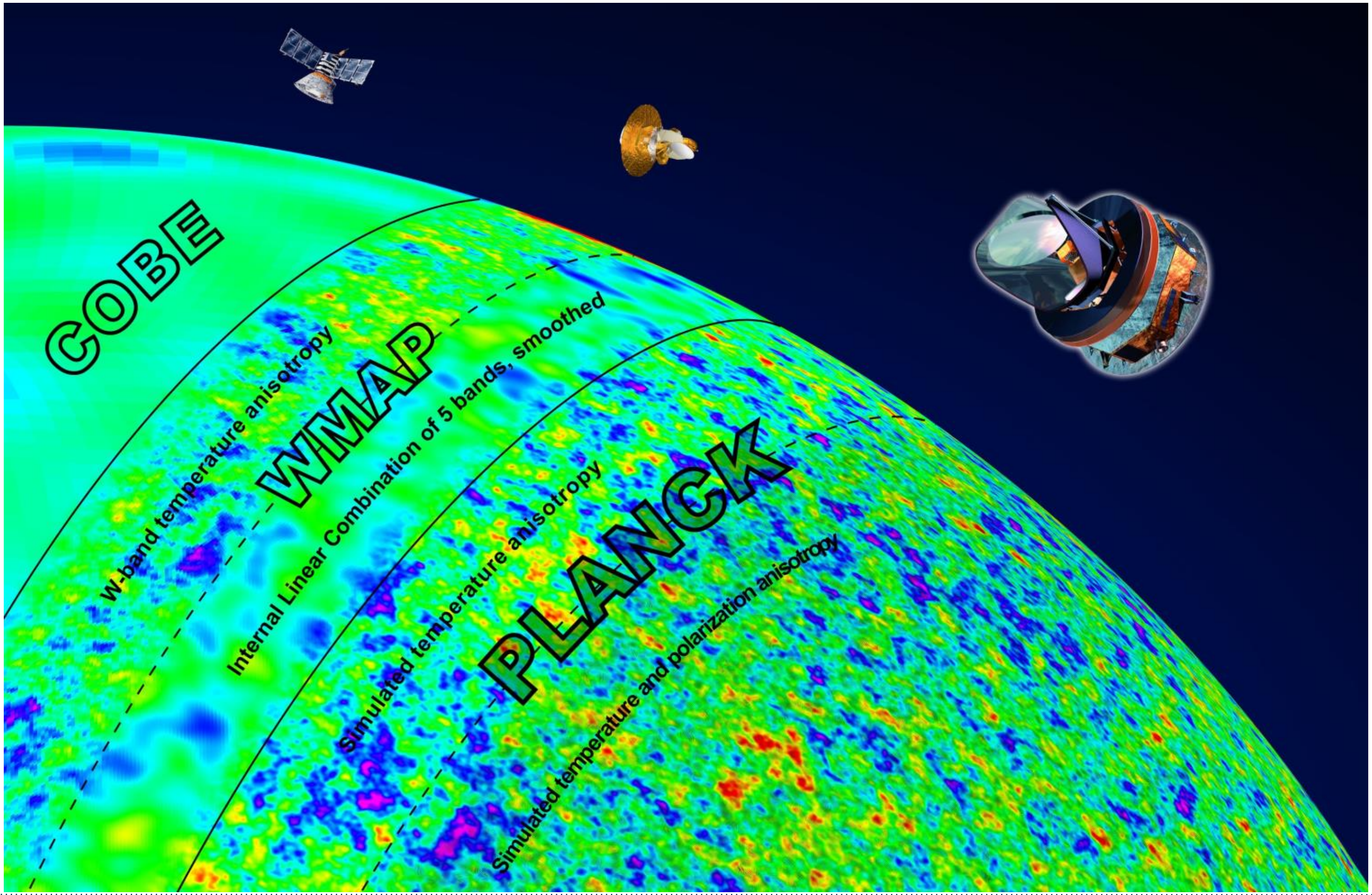


What Does The CMB Look Like?



CMB Science Evolution

Evolving science goals require (i) higher resolution & (ii) polarization sensitivity.



The CMB Data Challenge

- Extracting fainter signals (polarization, high resolution) from the data requires:
  - larger data volumes to provide higher signal-to-noise.
  - more complex analyses to control fainter systematic effects.

Experiment	Start Date	Observations	Pixels
COBE	1989	10^{9}	10^{4}
BOOMERanG	2000	10^{9}	10^{6}
WMAP	2001	10^{10}	10^{7}
Planck	2009	10^{12}	10^{9}
PolarBear	2012	10^{13}	10^{7}
QUIET-II	2015	10^{14}	10^{7}
CMBpol	2020+	10^{15}	10^{10}

- 1000x increase in data volume over last & next 15 years
  - need linear analysis algorithms to scale through next 10 M-foldings !

CMB Data Analysis

- In principle very simple
  - Assume Gaussianity and maximize the likelihood
    1. of maps given the observations and their noise statistics (analytic).
    2. of power spectra given maps and their noise statistics (iterative).

- In practice very complex
  - Correlated/colored noise
  - Non-deal data: foregrounds, glitches, asymmetric beams, etc.
  - Algorithm & implementation scaling with evolution of
    - CMB data-set size
    - HPC architecture

Analysis Algorithms

- Exact solutions involve both the map and its (dense) correlation matrix.
- Solutions scale as N_p^2 in memory, N_p^3 in operations - impractical for $N_p > 10^5$
- Require approximate solutions:
  - Solve for map only using preconditioned conjugate gradient
    - Scales as $N_i \, N_t$
  - Solve for pseudo-spectra only using spherical harmonic transforms
    - Scales as $N_p^{3/2}$
    - Biased by incomplete sky & inhomogeneous noise
  - Debias and quantify uncertainties using Monte Carlo methods: simulate and map $10^2 - 10^4$ realizations of the data
    - Scales as $N_r \, N_i \, N_t$

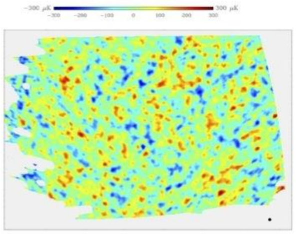
CMB Data Analysis Evolution

Data volume & computational capability dictate analysis approach.

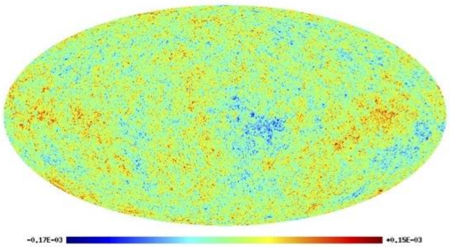
Date	Data	System	Map	Power Spectrum
1997 - 2000	B98	Cray T3E x 700	Explicit Maximum Likelihood (Matrix Invert - N_p^3)	Explicit Maximum Likelihood (Matrix Cholesky + Tri-solve - N_p^3)
2000 - 2003	B2K2	IBM SP3 x 3,000	Explicit Maximum Likelihood (Matrix Invert - N_p^3)	Explicit Maximum Likelihood (Matrix Invert + Multiply - N_p^3)
2003 - 2007	Planck SF	IBM SP3 x 6,000	PCG Maximum Likelihood (band-limited FFT – few N_t)	Monte Carlo (Sim + Map - many N_t)
2007 - 2010	Planck AF EBEX	Cray XT4 x 40,000	PCG Maximum Likelihood (band-limited FFT – few N_t)	Monte Carlo (SimMap - many N_t)
2010 - 2013	Planck MC PolarBear	Cray XE6 x 150,000	PCG Maximum Likelihood (band-limited FFT – few N_t)	Monte Carlo (Hybrid SimMap - many N_t)

Scaling In Practice

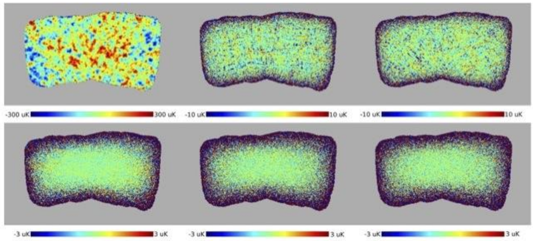
- 2000: BOOMERanG T-map
  - 10^8 samples => 10^5 pixels
  - 128 Cray T3E processors;



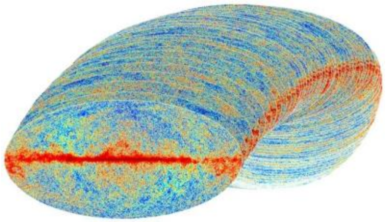
- 2006: Planck T-map
  - 10^{10} samples => 10^8 pixels
  - 6000 IBM SP3 processors;



- 2008: EBEX T/P-maps
  - 10^{11} samples => 10^6 pixels
  - 15360 Cray XT4 cores.



- 2010: Planck Monte Carlo 1000 T-maps
  - 10^{14} samples => 10^{11} pixels
  - 32000 Cray XT4 cores.



The Planck Challenge

- Most computationally challenging part of Planck analysis is simulating and mapping Monte Carlo realization sets.
- First Planck single-frequency simulation & map-making took 6 hours on 6000 CPUs (36,000 CPU-hours per realization) in 2006.
- Our goal was 10,000 realizations of all 9 frequencies in 2012
  – With no change => 3×10^9 CPU-hours
  – With Moore's Law => 2×10^8 CPU-hours
  – NERSC quota => $O(10^7)$ CPU-hours
- Required
  – Ability to scale through 4 epochs of Moore's Law, however they might be realized (clock speed, concurrency, accelerators, ?)
  – Additional O(20x) algorithmic/implementation speed-up

Simulation & Mapping: Calculations

Given the instrument noise statistics & beams, a scanning strategy, and a sky:

1) SIMULATION: $d_t = n_t + s_t = n_t + P_{tp}\, s_p$
   - A realization of the piecewise stationary noise time-stream:
     - Pseudo-random number generation & FFT
   - A signal time-stream scanned & beam-smoothed from the sky map:
     - SHT

2) MAPPING: $(P^T N^{-1} P)\, d_p = P^T N^{-1} d_t$ $\quad$ $(A\, x = b)$
   - Build the RHS
     - FFT & sparse matrix-vector multiply
   - Solve for the map
     - PCG over FFT & sparse matrix-vector multiply

Simulation & Mapping: Scaling

- In theory such analyses should scale
  - Linearly with the number of observations.
  - Perfectly to arbitrary numbers of cores.
- In practice this does not happen because of
  - IO (reading pointing; writing time-streams
    reading pointing & timestreams; writing maps)
  - Communication (gathering maps from all processes)
  - Calculation inefficiency (linear operations => minimal data re-use)
- Code development has been an ongoing history of addressing these challenges anew with each new data volume and system concurrency.

IO - Before

For each MC realization

For each detector

Read detector pointing

Write detector timestream

For all detectors

Read detector timestream & pointing

Write map

Sim

Map

$\Rightarrow$ Read: Realizations x Detectors x Observations x 2

Write: Realizations x (Detectors x Observations + Pixels)

E.g. for Planck read 500PB & write 70PB.

IO - Optimizations

- Read sparse telescope pointing instead of dense detector pointing
  – Calculate individual detector pointing on the fly.

- Remove redundant write/read of time-streams between simulation & mapping
  – Generate simulations on the fly only when map-maker requests data.

- Put MC loop inside map-maker
  – Amortize common data reads over all realizations.

IO – After

Read telescope pointing
For each detector
    Calculate detector pointing
For each MC realization
    For all detectors
        Simulate time-stream
    Write map

SimMap

$\Rightarrow$ Read: Sparse Observations
Write: Realizations x Pixels

E.g. for Planck, read 2GB & write 70TB => 10^8 read & 10^3 write compression.

Communication Details

- The time-ordered data from all the detectors are distributed over the processes subject to:
  - Load-balance
  - Common telescope pointing
- Each process therefore holds
  - *some* of the observations
  - for *some* of the pixels.
- In each PCG iteration, each process solves with its observations.
- At the end of each iteration, each process needs to gather the total result for all of the pixels in its subset of the observations.

Communication - Before

- Initialize a process & MPI task on every core
- Distribute time-stream data & hence pixels
- After each PCG iteration
  - Each process creates a full map vector by zero-padding
  - Call MPI_Allreduce(map, world)
  - Each process extracts the pixels of interest to it & discards the rest

Communication – Optimizations

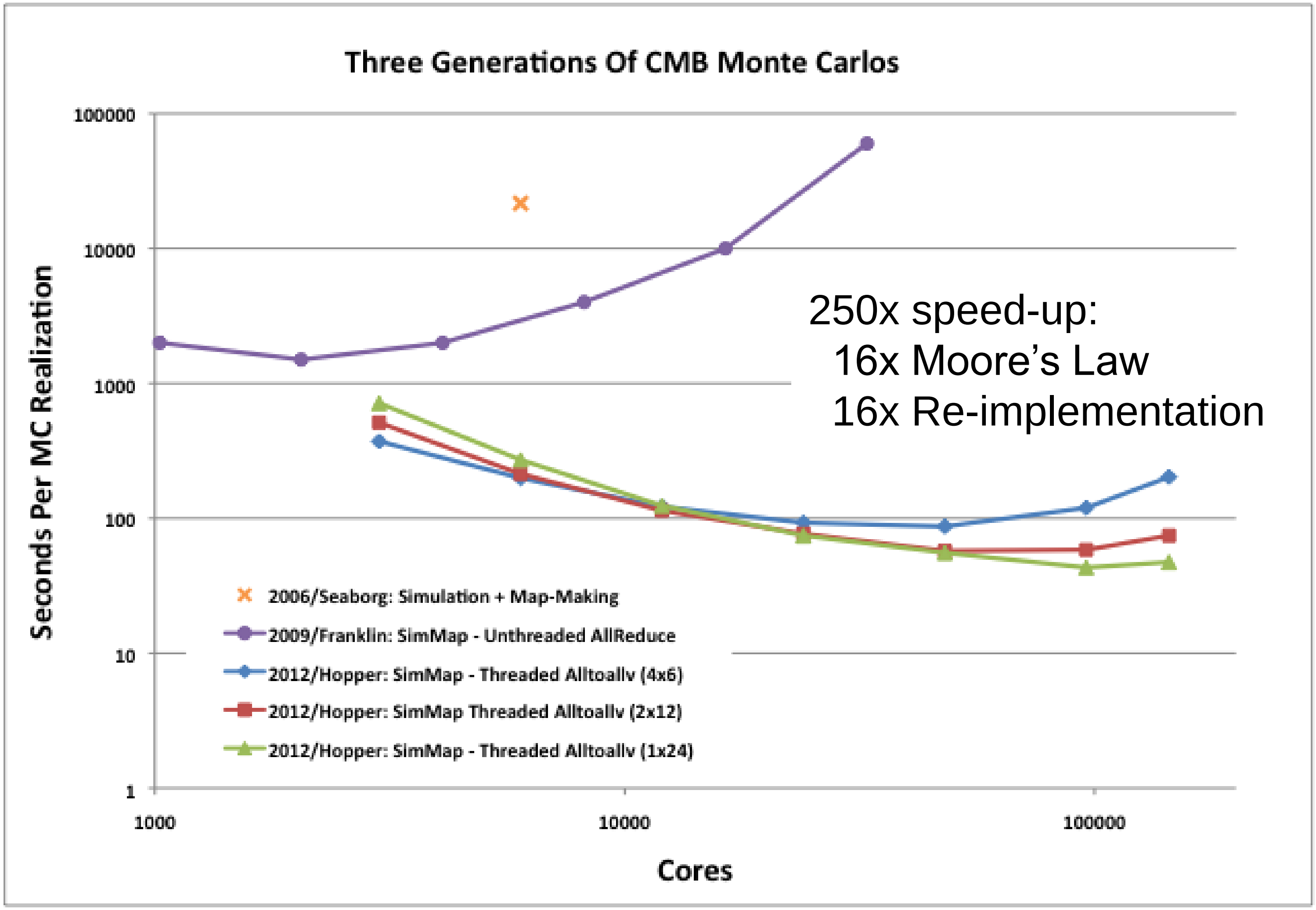
- Reduce the number of MPI tasks
  - Only use MPI for off-node communication
  - Use threads on-node

- Minimize the total volume of the messages
  - Determine all processes' pair-wise pixel overlap
  - If the data volume is smaller, use *scatter/gather* in place of *reduce*

Communication – ~~After~~ Now

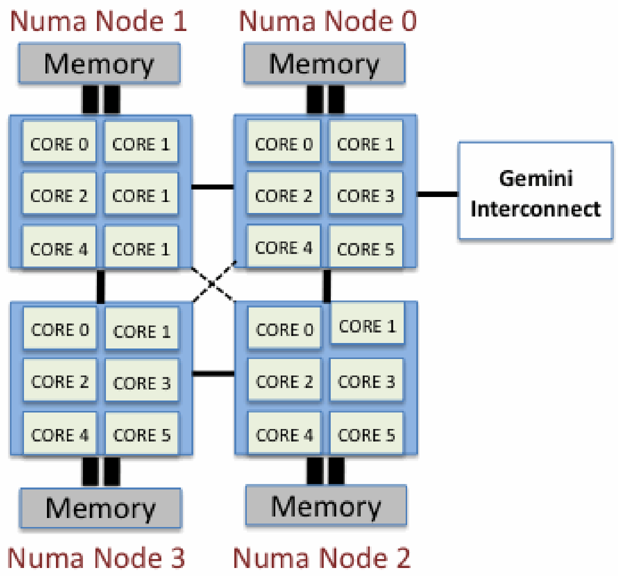
- Initialize a process & MPI task on every node
- Distribute time-stream data & hence pixels
- Calculate common pixels for every pair of processes
- After each PCG iteration
  - If most pixels are common to most processes
    - use MPI_Allreduce(map, world) as before
  - Else
    - Each process prepares its send buffer
    - Call MPI_Alltoallv(sbuffer, rbuffer, world)
    - Each process only receives/accumulates data for pixels it sees.

Planck Simulations Over Time

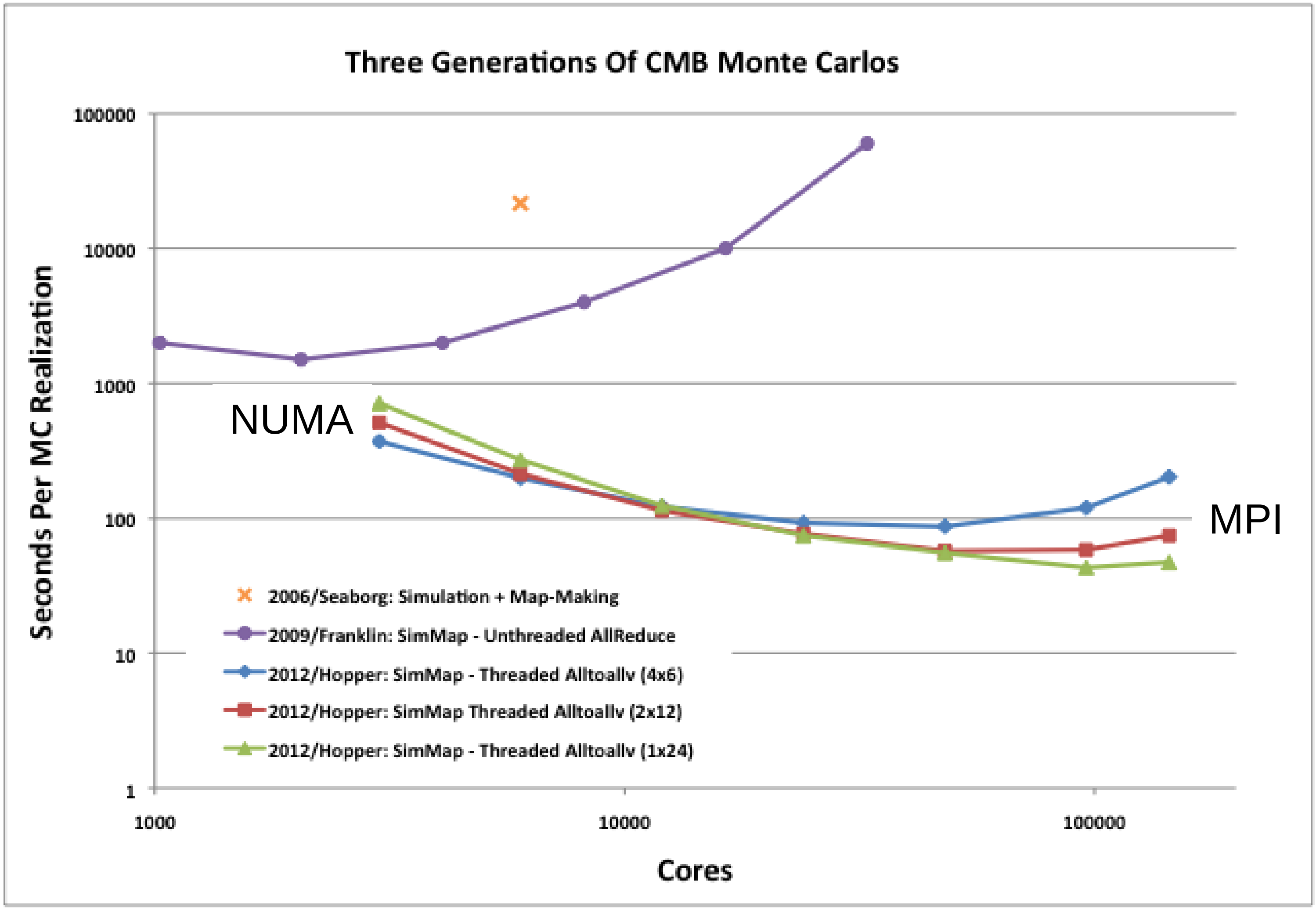


HPC System Evolution

- Clock speed is no longer able to maintain Moore's Law.
- Multi-core CPU and GPGPU are two major approaches.
- Both of these will require
  - significant code development
  - performance experiments & auto-tuning
- E.g. NERSC's Cray XE6 system *Hopper*
  - 6384 nodes
  - 2 sockets per node
  - 2 NUMA nodes per socket
  - 6 cores per NUMA node
- What is the best way to run hybrid code on such a system?

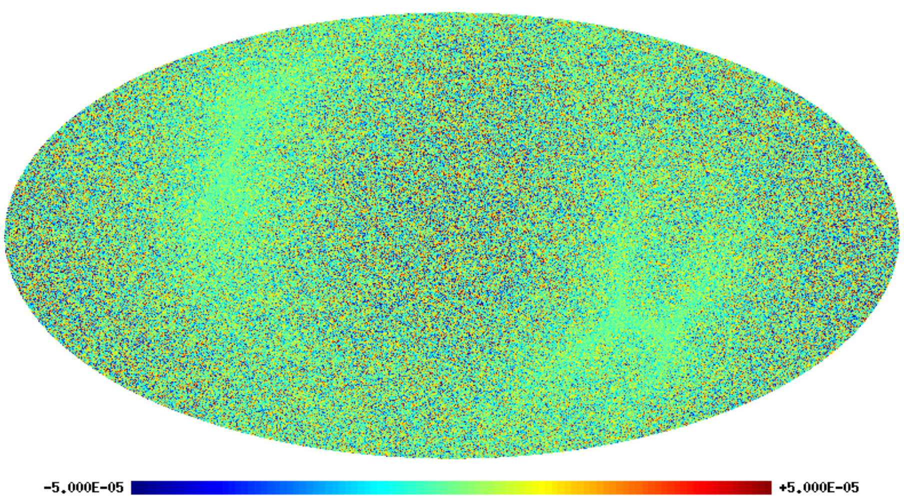


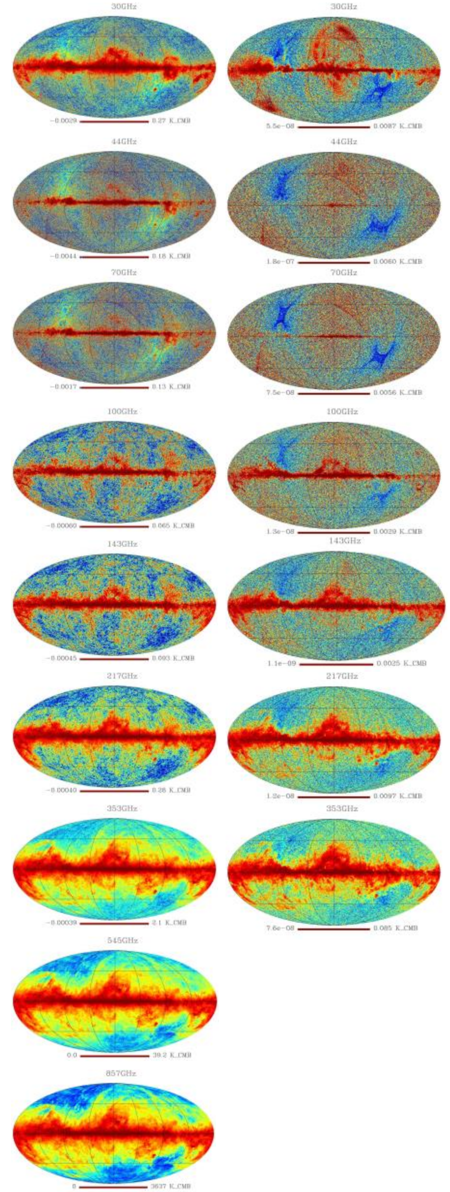
Configuration With Concurrency



Planck Full Focal Plane 6

- 6th full-mission simulation set - key to 2013 results.
- Single fiducial sky for validation & verification.
- 1,000 CMB & noise realizations for debiasing and uncertainty quantification.
- 250,000 maps in total – largest CMB MC set ever.
- 2014 & 15 releases will require 10,000 realizations.





Planck March 2013 Results

- 28 papers released by the collaboration
- Cosmology highlights
  - Data very well fit by 6 parameter model
  - Some tension with previous results
    - 2% more dark matter, less dark energy
    - 10% lower Hubble constant (2.5σ)
  - Map of all dark matter via lensing
  - 3 light neutrino species ($\Sigma m < 0.23$eV)
  - Scalar/tensor ratio $r < 0.1$
  - Possible asymmetry & outliers
  - All results tested against FFP6


The New York Times

PRESIDENT URGES ISRAELIS TO PUSH EFFORT FOR PEACE

The Cosmos, Back in the Day

Bronx Inspector, Secretly Taped, Suggests Race Is a Factor in Stops

Once Few, Women Hold More Power in Senate

Fast-Growing Brokerage Firm Often Tangles With Regulators

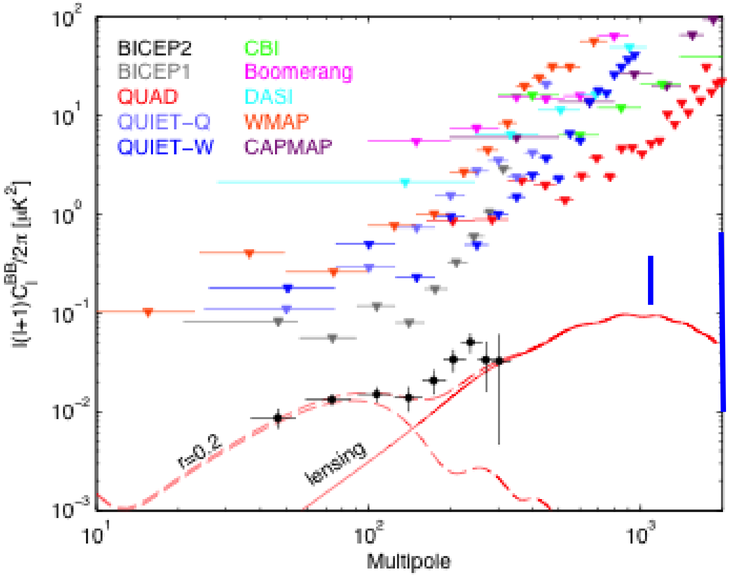
O Revelations! Letters, Once Banned, Flesh Out Willa Cather

COMPUTATIONAL COSMOLOGY CENTER

BERKELEY LAB

BICEP2 Results

- BICEP2 recently announced r ~ 0.2
  - Much higher than expected; inconsistent* with Planck
  - Predicted to result in 3rd Nobel prize for CMB work
- Many reasons for skepticism – await Planck results later this year
  - http://www.facebook.com/groups/574544055974988/



Future Prospects

- Next-generation B-mode experiments will gather
  - 10x Planck: current suborbital
  - 100x Planck: future suborbital
  - 1000x Planck: future satellite (or multi-site suborbital)

- Next-generation supercomputers will have
  - Huge core counts
  - Increasingly heterogeneous nodes
  - Varied accelerators (GPGPU, MIC, ?, ?)
  - Increasingly constrained power

Conclusions

- The CMB provides a unique window onto the early Universe
  - investigate fundamental cosmology & physics.
- CMB data analysis is a computationally-challenging problem requiring state of the art HPC capabilities.
- Both the CMB data sets we are gathering and the HPC systems we are using to analyze them are evolving – this is a persistent, dynamic problem.
- The science we can extract from present and future CMB data sets will be determined by the limits on
  a) our computational capability, and
  b) our ability to exploit it.